

Grau de dependência da Renda média com a variável liberdade econômica e inovação

A modelagem que será apresentada buscou estudar o quanto as variáveis inovação e liberdade econômica influenciam o nível de renda per capita de um país. Pegou-se os dados de renda per capita de 112 países, além de seus respectivos índices de inovação e liberdade econômica, para realizar essa modelagem estatística.

Antes de tudo é bom explicar o que é uma regressão. Regressão linear nada mais é que uma forma de modelagem estatística que visa estudar a relação de dependência de uma variável (Y, ou variável dependente) com uma ou mais variáveis (x's, ou variáveis explicativas). Explicando de uma forma simples e um pouco grosseira, esse tipo de modelagem tenta entender o quanto, na média, a variável explicativa pode influenciar ou explicar a variável dependente.

No caso desse estudo, tentou-se entender a dependência da variável PIB per capita, Y_i , com as variáveis inovação, xI , e liberdade econômica, xL . Primeiro, rodou-se dois modelos simples: o primeiro colocando $Y_i \sim xL$, ou seja, o quanto o pib per capita era explicado pela variável liberdade econômica. O segundo modelo era $Y_i \sim xI$, portanto, o quanto a inovação explica a variável pib per capita. Por último, tenta-se entender o quanto as variáveis inovação e liberdade econômica combinadas explicam o pib per capita.

O estudo visou apenas entender se a inovação e a liberdade econômicas tendem ou não a gerar efeitos positivos ou negativos para o nível de renda de um país.

Limitações do modelo:

1. O grau de desenvolvimento econômico e algumas outras características (como crescimento, cenário macro e microeconômico e ainda outras coisas não captadas por esse estudo inicial) influenciam o grau de inovação tecnológico de um país, segundo vários e vários estudos. Por exemplo, países como Brasil, China e Índia, que são países de renda média têm desempenhos na área de inovação acima do esperado. O grau de desenvolvimento, no entanto, é um fenômeno construído ao longo do tempo.
2. A modelagem não comporta o ponto 1, uma vez que ela se utiliza de dados apenas do ano passado e não de uma série temporal. Estudos posteriores deveriam pegar de uma longa série temporal para obter resultados mais precisos. Isto é, dever-se-ia realizar

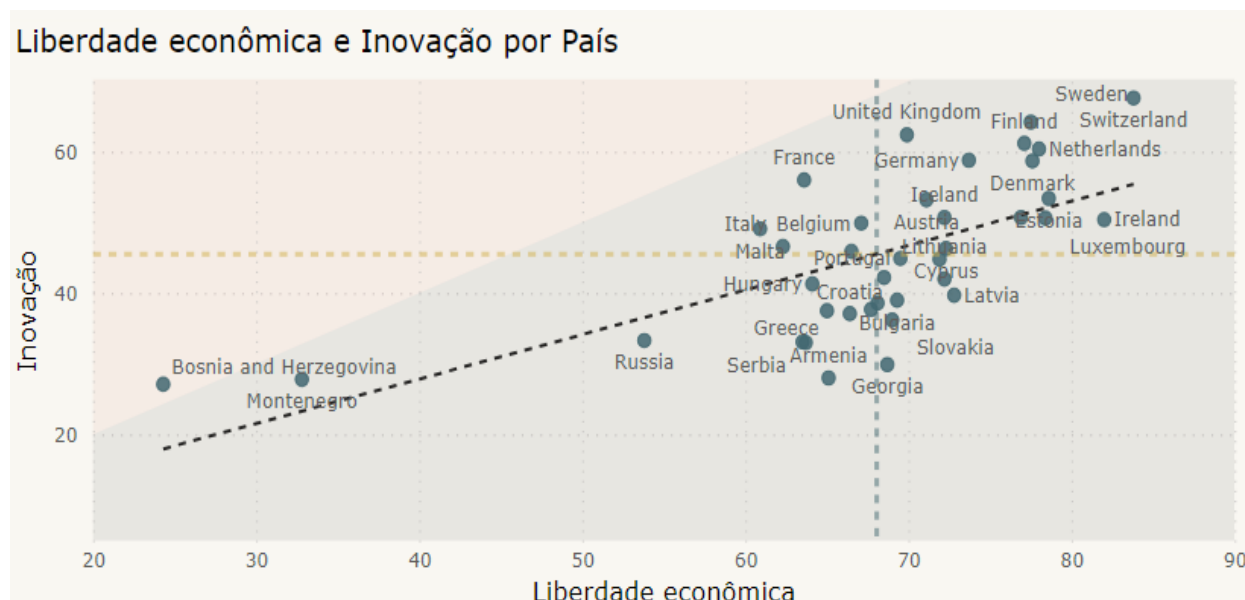
essa modelagem ao longo do tempo e ver se a inovação e a liberdade econômica influenciam a renda positiva ou negativamente ao longo desse período na média

3. O modelo de regressão linear possui alguns pressupostos. Desses pressupostos, o que preocupou mais foi o da multicolinearidade, isto é, a possível alta correlação entre as variáveis inovação e liberdade econômica. Entretanto, depois do teste para se verificar isso chegou-se à conclusão da não existência de multicolinearidade relevante. Os demais pressupostos parecem estarem também ok's também
4. O índice de inovação faz o coeficiente de determinação explodir! Ou seja, com base nesse estudo a inovação explica muito da renda per capita de um país. Isso pode ser outra falha desse estudo. Posteriormente, dever-se-ia realizar alguns procedimentos, como aumentar a amostra, para se verificar a consistência desse fato observado

Resultados da modelagem e a interpretação:

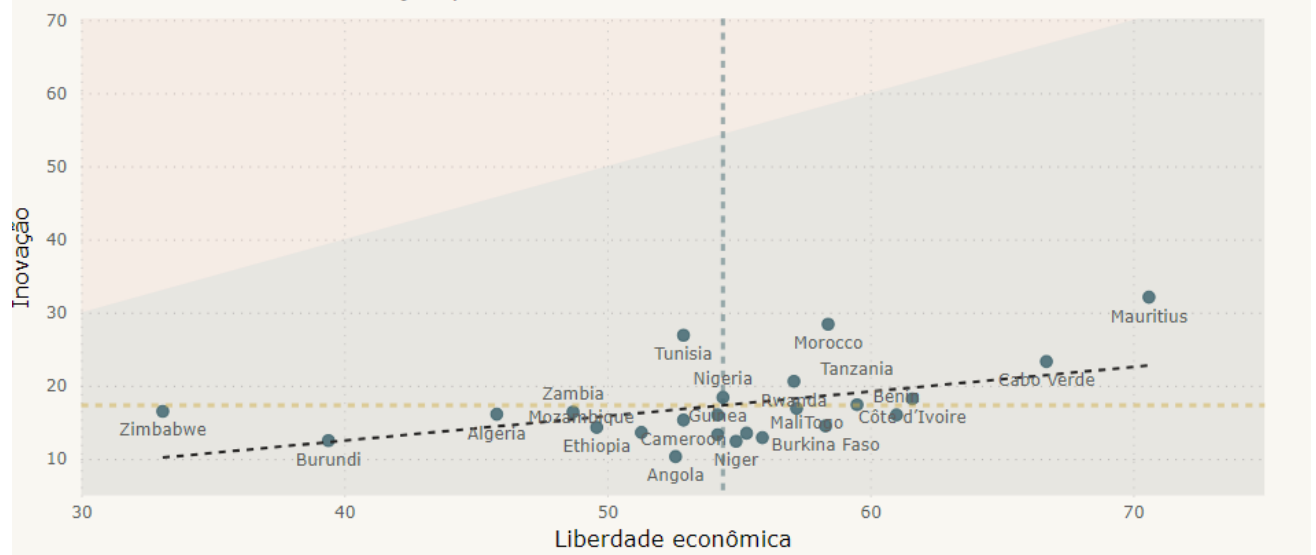
- O grau de inovação de um país parece explicar mais a renda per capita que a liberdade econômica é capaz;
- Os países mais inovadores tendem a ter uma maior liberdade econômica

Europa:



África:

Liberdade econômica e Inovação por País

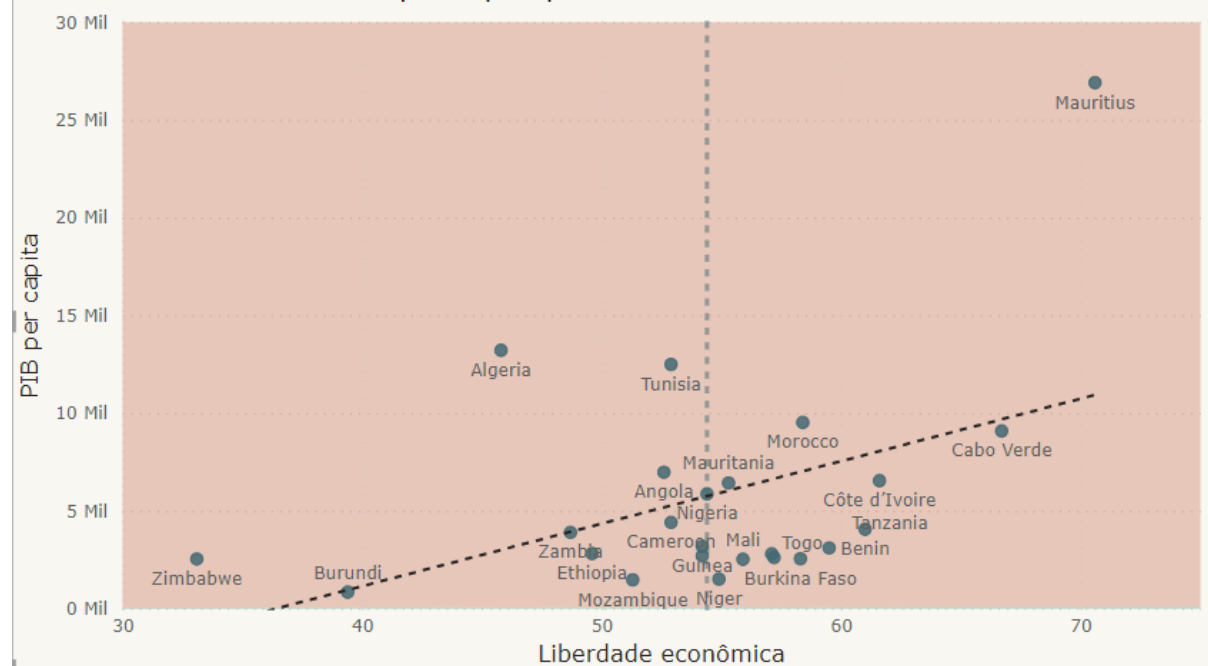


Feito no BI 😊!

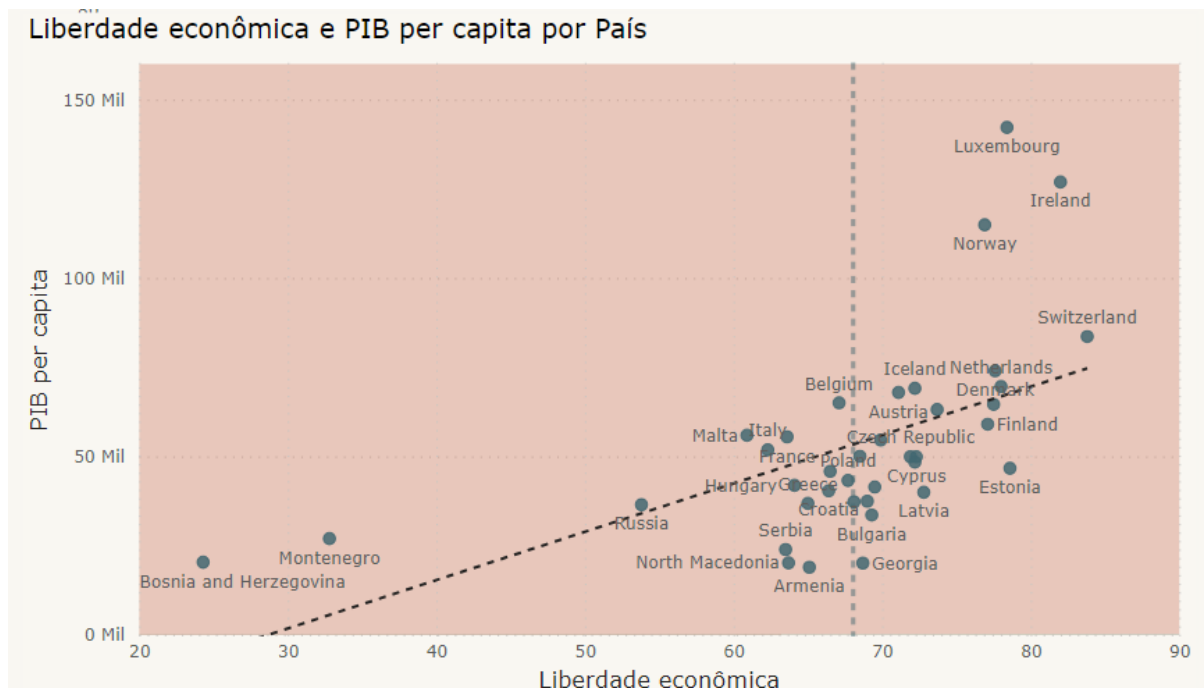
- Países que se utilizam da combinação inovação + liberdade econômica têm renda per capita muito maiores na média:

África:

Liberdade econômica e PIB per capita por País



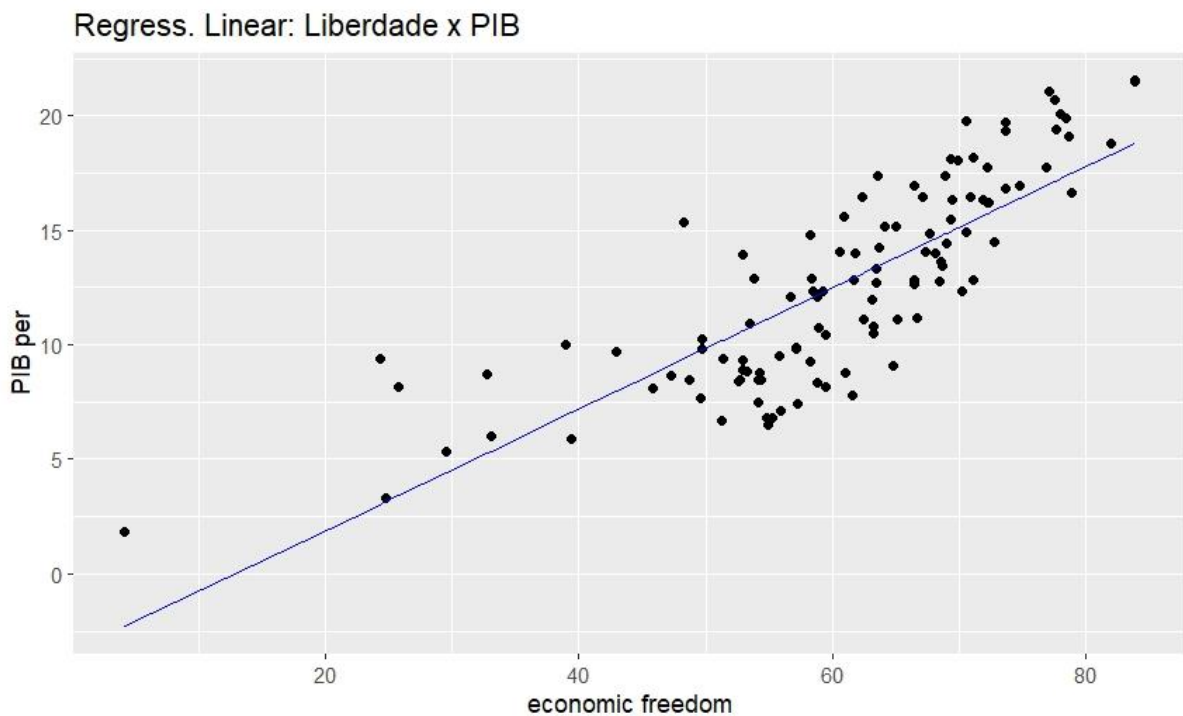
Europa:



Todos os resultados foram obtidos via R.

Nele vemos, em resumo, que o modelo de regressão parece adequado, pois as variáveis independentes x_I e x_L são estatisticamente significativas e o modelo explica uma proporção alta da variabilidade em Y . Não se deve esquecer as limitações apontadas.

Regressão liberdade econômica vs. Renda per capita:



`summary(regL)` --> É o modelo $Y \sim xL$

Call:

`lm(formula = Y ~ xL, data = dados)`

Residuals:

Min 1Q Median 3Q Max

-5.0879 -1.8909 0.0256 2.0078 6.3575

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -3.40178 1.12739 -3.017 0.00317 **

xL 0.26474 0.01819 14.555 < 2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.562 on 110 degrees of freedom

Multiple R-squared: 0.6582, Adjusted R-squared: 0.6551

F-statistic: 211.9 on 1 and 110 DF, p-value: $< 2.2e-16$

•Interpretação dos resultados: Esta saída representa os resultados de uma regressão linear simples,

onde Y é a variável dependente e xL é a variável independente. Aqui está a interpretação dos resultados:

Coefficientes:

O coeficiente para a variável xL é 0.26474. Isso significa que, em média, para cada aumento de

uma unidade em xL, Y aumenta em 0.26474 unidades, mantendo todas as outras variáveis constantes.

O coeficiente para a interceptação (Intercepto) é -3.40178. Isso significa que quando xL é zero,

espera-se que Y seja aproximadamente -3.40178.

Valores-p:

O valor-p para xL é $< 2e-16$, indicando que o coeficiente para xL é estatisticamente significativo.

O valor-p para a interceptação é 0.00317, indicando que o intercepto também é estatisticamente significativo.

Estatísticas de ajuste:

O R^2 é 0.6582, o que significa que aproximadamente 65.82% da variabilidade em Y é explicada pela variável xL.

O R^2 ajustado é 0.6551, que leva em consideração o número de variáveis independentes no modelo.

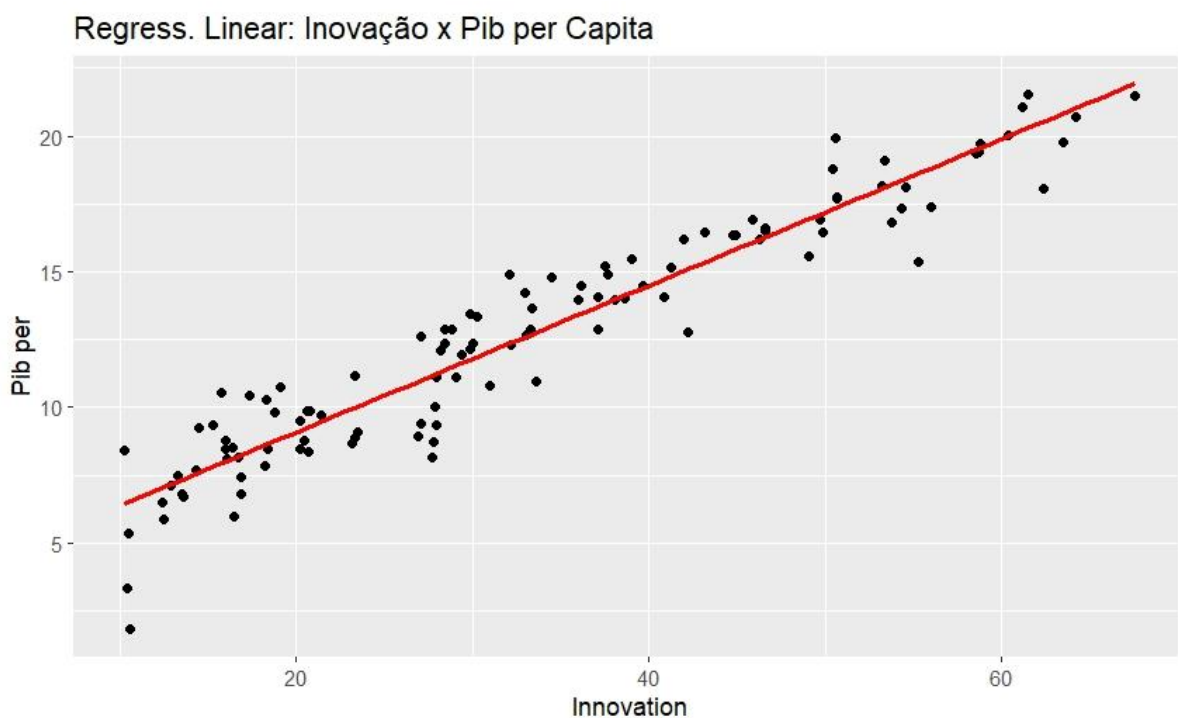
Resíduos: A média dos resíduos é próxima de zero, indicando que o modelo não apresenta um viés sistemático.

Os resíduos variam entre -5.0879 e 6.3575, o que sugere uma distribuição razoavelmente simétrica.

Erro padrão residual: O erro padrão residual é 2.562, indicando a dispersão dos pontos em torno da linha de regressão.

Em resumo, o modelo de regressão parece adequado, pois a variável independente xL é estatisticamente significativa e o modelo explica uma proporção razoavelmente alta da variabilidade em Y.

Regressão inovação econômica vs. Renda per capita:



```
summary(regI) -->
```

Call:

```
lm(formula = Y ~ xI, data = dados)
```

Residuals:

Min 1Q Median 3Q Max

-4.7211 -0.7733 0.1250 0.9556 2.5768

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 3.665766 0.299269 12.25 <2e-16 ***

xI 0.270325 0.008197 32.98 <2e-16 ***

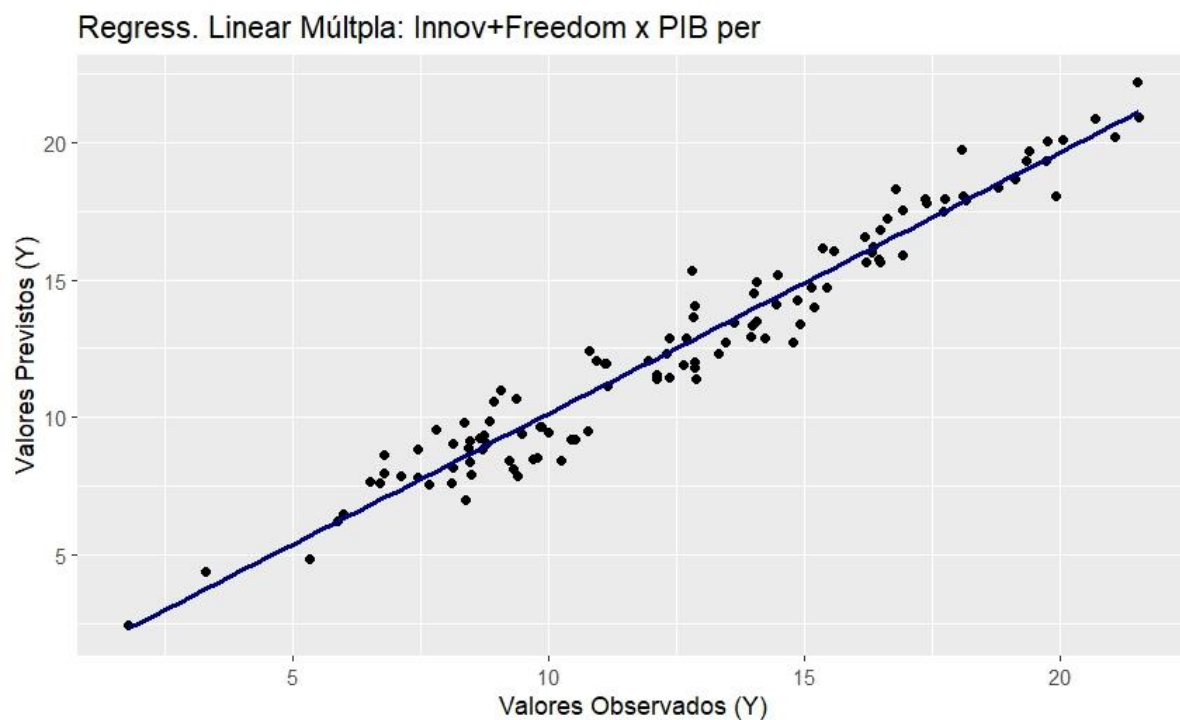
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.328 on 110 degrees of freedom

Multiple R-squared: 0.9081, Adjusted R-squared: 0.9073

F-statistic: 1087 on 1 and 110 DF, p-value: < 2.2e-16

Regressão liberdade econômica + inovação vs. Renda per capita:



summary(reg)

Call:

lm(formula = Y ~ xI + xL, data = dados)

Residuals:

Min 1Q Median 3Q Max

-2.52417 -0.63680 0.04152 0.69827 2.08063

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -0.219168 0.437127 -0.501 0.617

xI 0.212942 0.008141 26.156 <2e-16 ***

xL 0.095590 0.009365 10.207 <2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.954 on 109 degrees of freedom

Multiple R-squared: 0.953, Adjusted R-squared: 0.9522

F-statistic: 1106 on 2 and 109 DF, p-value: < 2.2e-16

- Interpretação do resultados:

Esta saída representa os resultados de um modelo de regressão linear múltipla, onde Y é a variável dependente e xI e xL são as variáveis independentes. Aqui está a interpretação dos resultados:

Coeficientes: Para xI, o coeficiente é 0.212942. Isso significa que, mantendo xL constante, um aumento de uma unidade em xI está associado a um aumento de 0.212942 unidades em Y.

Para xL, o coeficiente é 0.095590. Isso significa que, mantendo xI constante, um aumento de uma unidade em xL está associado a um aumento de 0.095590 unidades em Y.

O intercepto é -0.219168, que é o valor esperado de Y quando xI e xL são zero.

Valores-p: Ambos os coeficientes para xI e xL têm valores-p muito baixos (< 2e-16), indicando que ambos são estatisticamente significativos.

O valor-p para o intercepto é 0.617, o que sugere que o intercepto não é estatisticamente significativo.

Estatísticas de ajuste: O R^2 é 0.953, o que significa que aproximadamente 95.3% da variabilidade em Y é explicada pelas variáveis xI e xL neste modelo.

O R^2 ajustado é 0.9522, que leva em consideração o número de variáveis independentes no modelo. Resíduos: A média dos resíduos é próxima de zero, indicando que o modelo não apresenta um viés sistemático. Os resíduos variam entre -2.52417 e 2.08063, o que sugere uma distribuição razoavelmente simétrica.

Erro padrão residual: O erro padrão residual é 0.954, indicando a dispersão dos pontos em torno da linha de regressão.